



HAL
open science

PERFORM PROTOCOL TESTS AND SEQUENCE CHECKS ? THE EXPERIENCE FROM MULTIOMICS MICROBIOTA DATA

Maria Bernard, Jean-Luc Coville, Nicolas Bruneau, Deborah Jardet, Fanny Calenge, Géraldine Pascal, Tatiana Zerjal

► To cite this version:

Maria Bernard, Jean-Luc Coville, Nicolas Bruneau, Deborah Jardet, Fanny Calenge, et al.. PERFORM PROTOCOL TESTS AND SEQUENCE CHECKS ? THE EXPERIENCE FROM MULTIOMICS MICROBIOTA DATA. 25ème Séminaire des Doctorants du Département de Génétique Animale (SDDGA 2022), Sep 2022, Bordeaux, France. . hal-03930031

HAL Id: hal-03930031

<https://hal.science/hal-03930031>

Submitted on 9 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

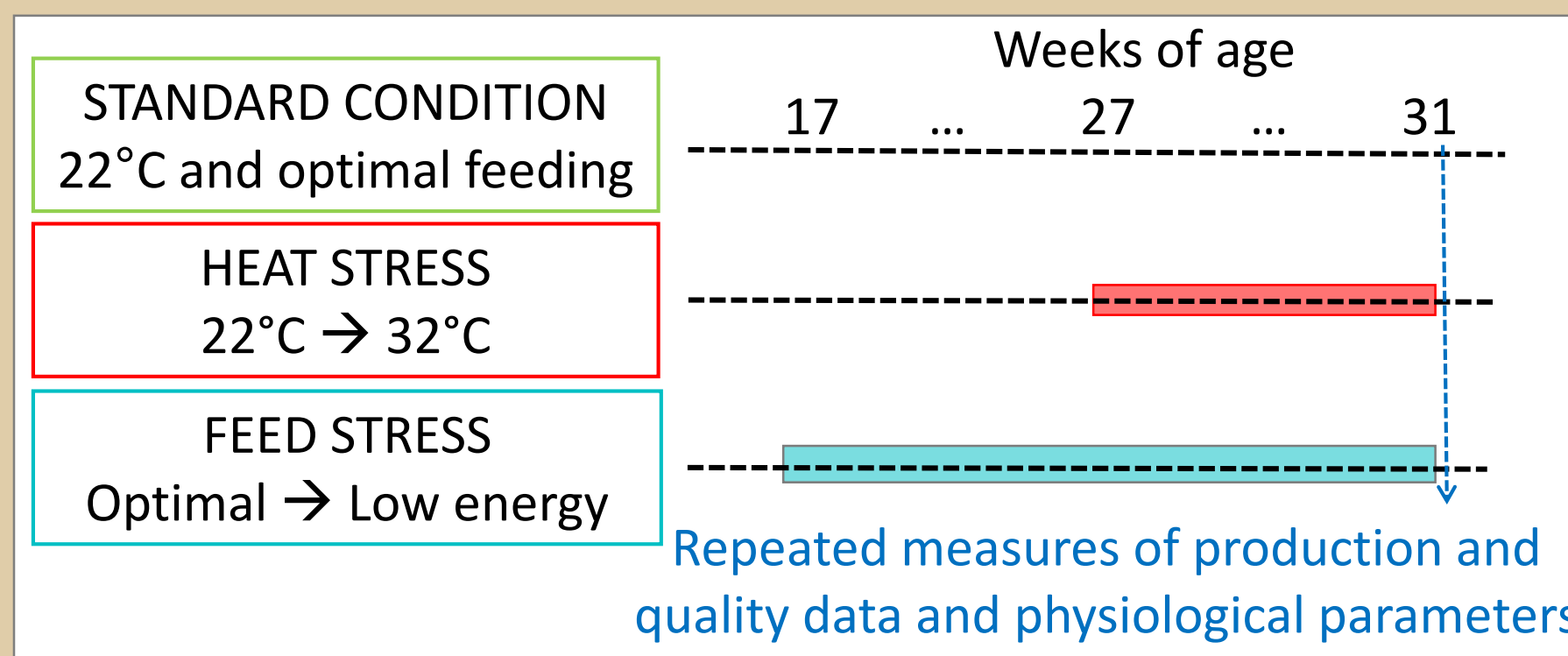
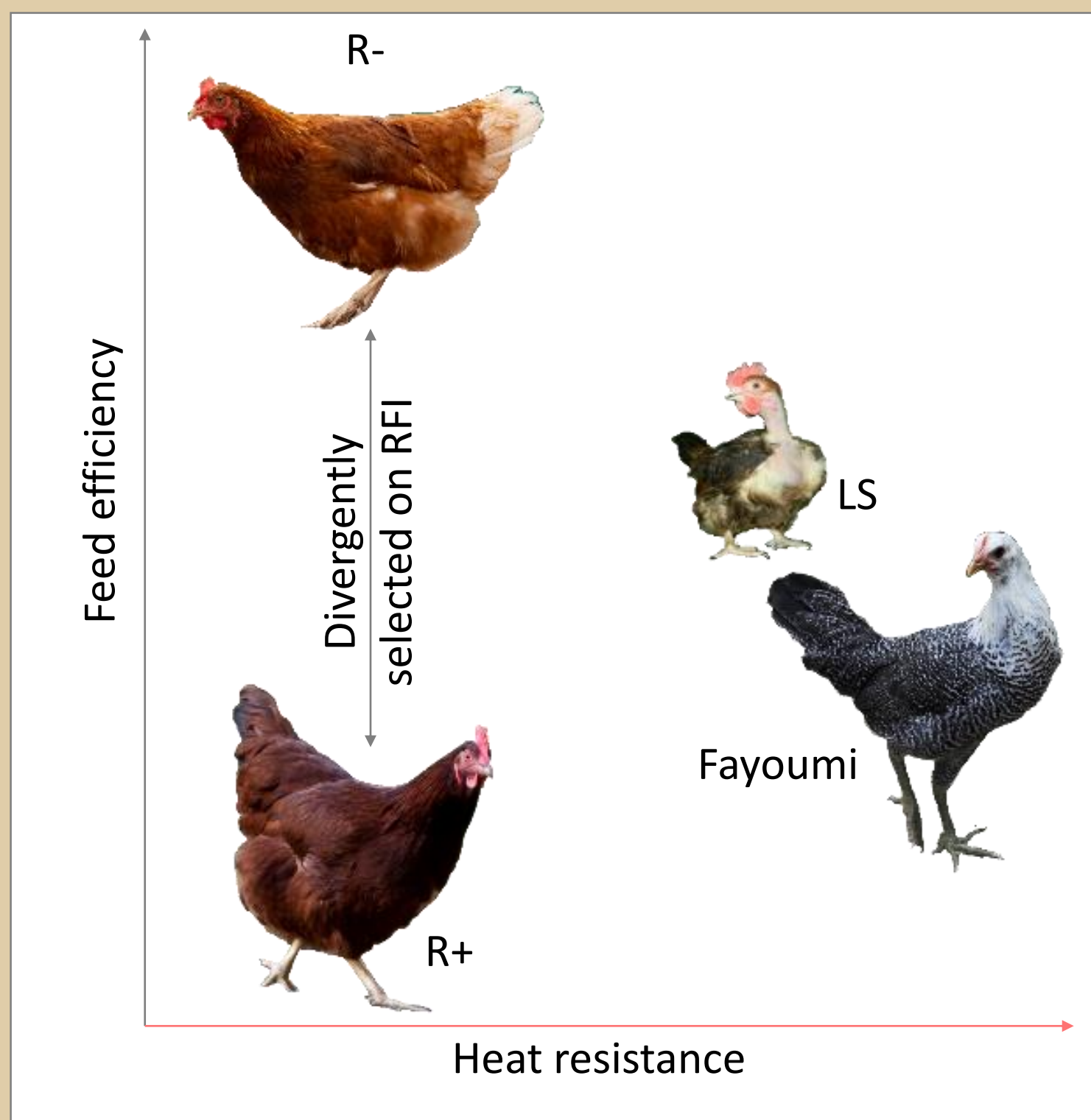
WHY PERFORM PROTOCOL TESTS AND SEQUENCE CHECKS ? THE EXPERIENCE FROM MULTI-OMICS MICROBIOTA DATA

Maria Bernard¹, Jean-Luc Coville¹, Nicolas Bruneau¹, Déborah Jardet¹, Fanny Calenge¹, Géraldine Pascal², Tatiana Zerjal¹

¹ : Université Paris-Saclay, INRAE, AgroParisTech, GABI, 78350 Jouy-en-Josas, France; ² : GenPhySE, Université de Toulouse, INRAE, ENVT, F-31326 Castanet Tolosan, France

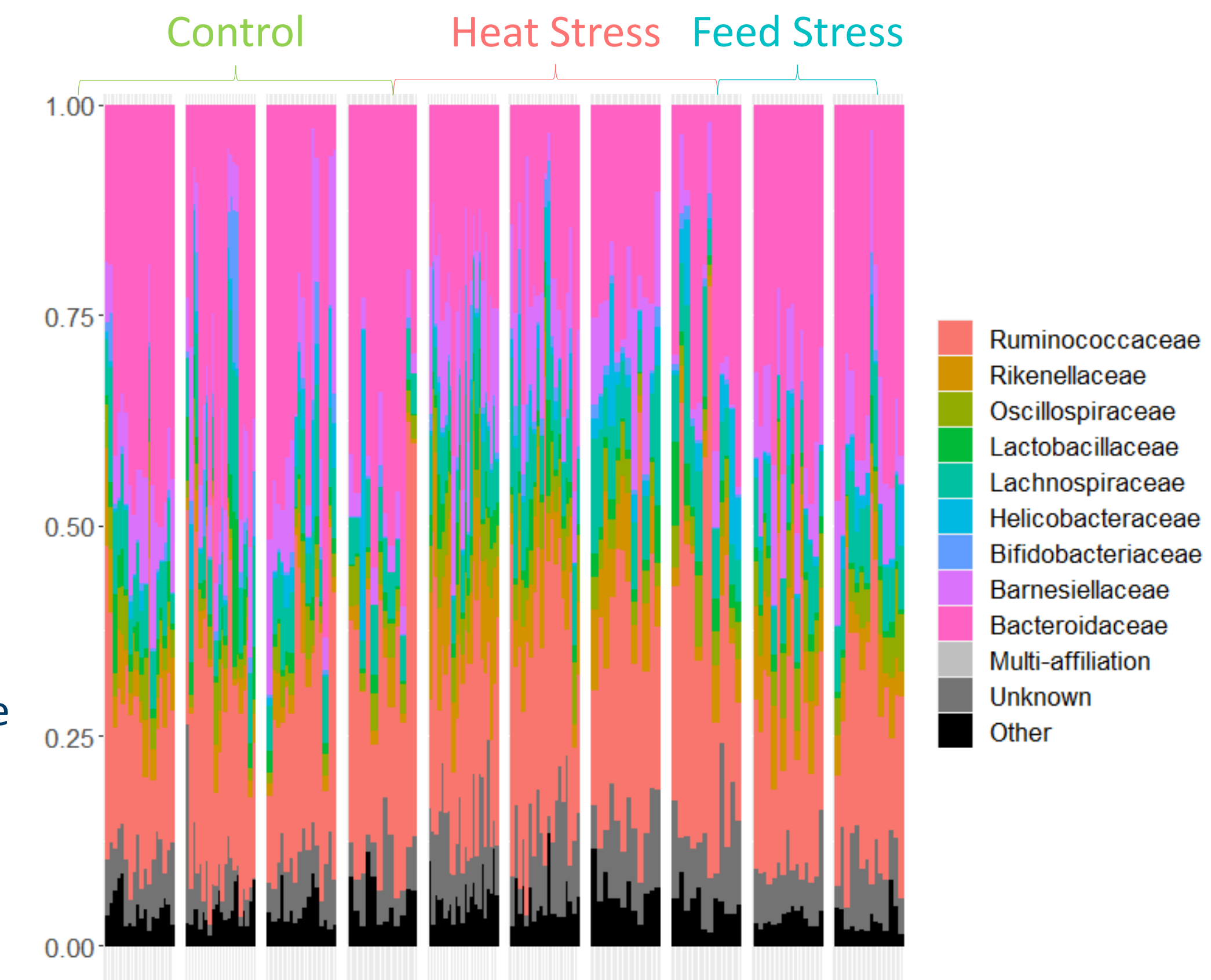
How to differentiate high microbiota natural variability from technical bias?

EXPERIMENTAL DESIGN



16S sequences have been analysed with FROGS (1), and bacterial composition characterised with Phyloseq (2) R package through:

- Family composition profiles
- 4 α -diversity indices
- 3 β -dissimilarity methods used for MDS and hierarchical clustering



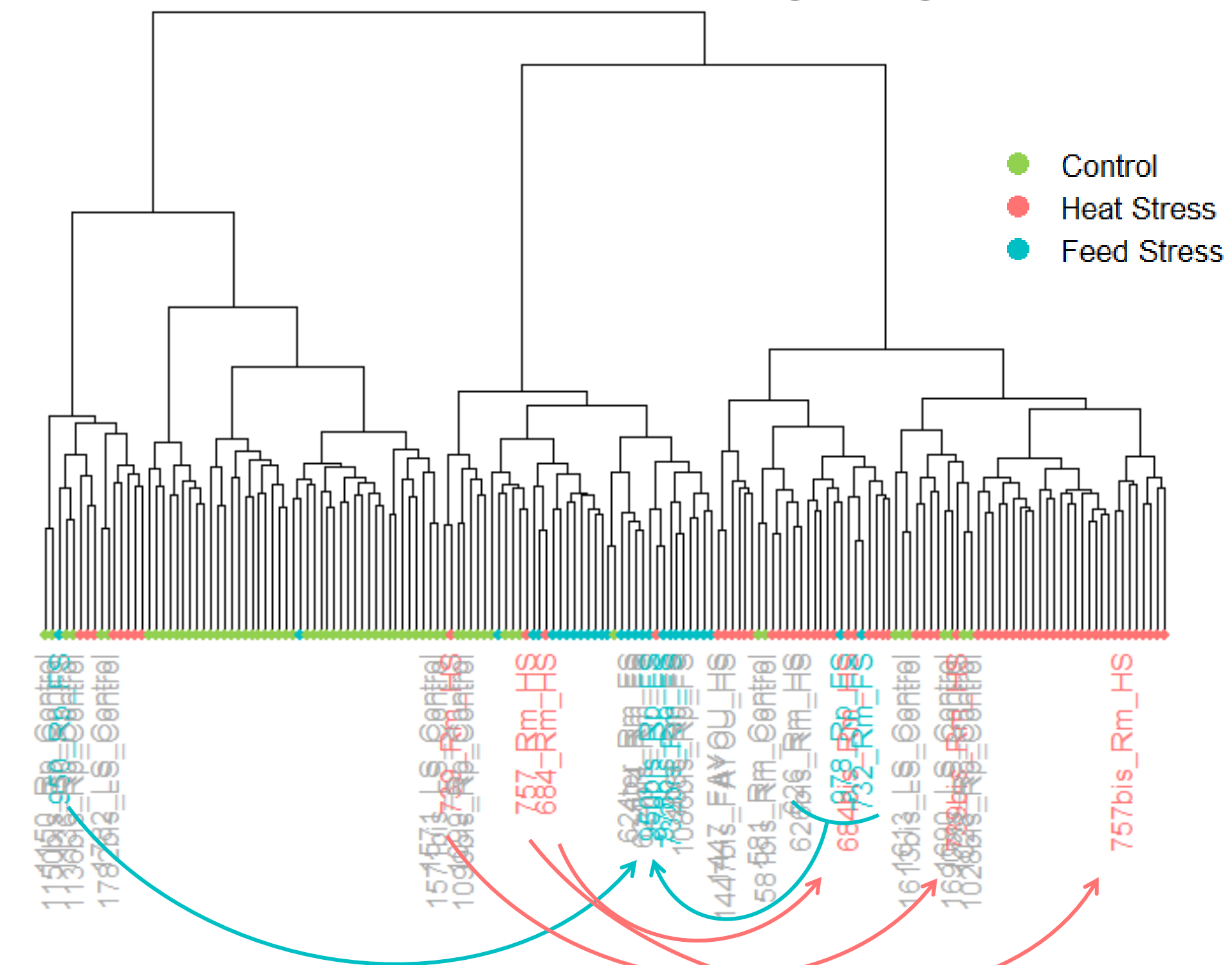
➔ Should the 14 suspicious samples be considered as samples with high natural variability or outliers ?

Resequencing of 20 samples allowed us to:

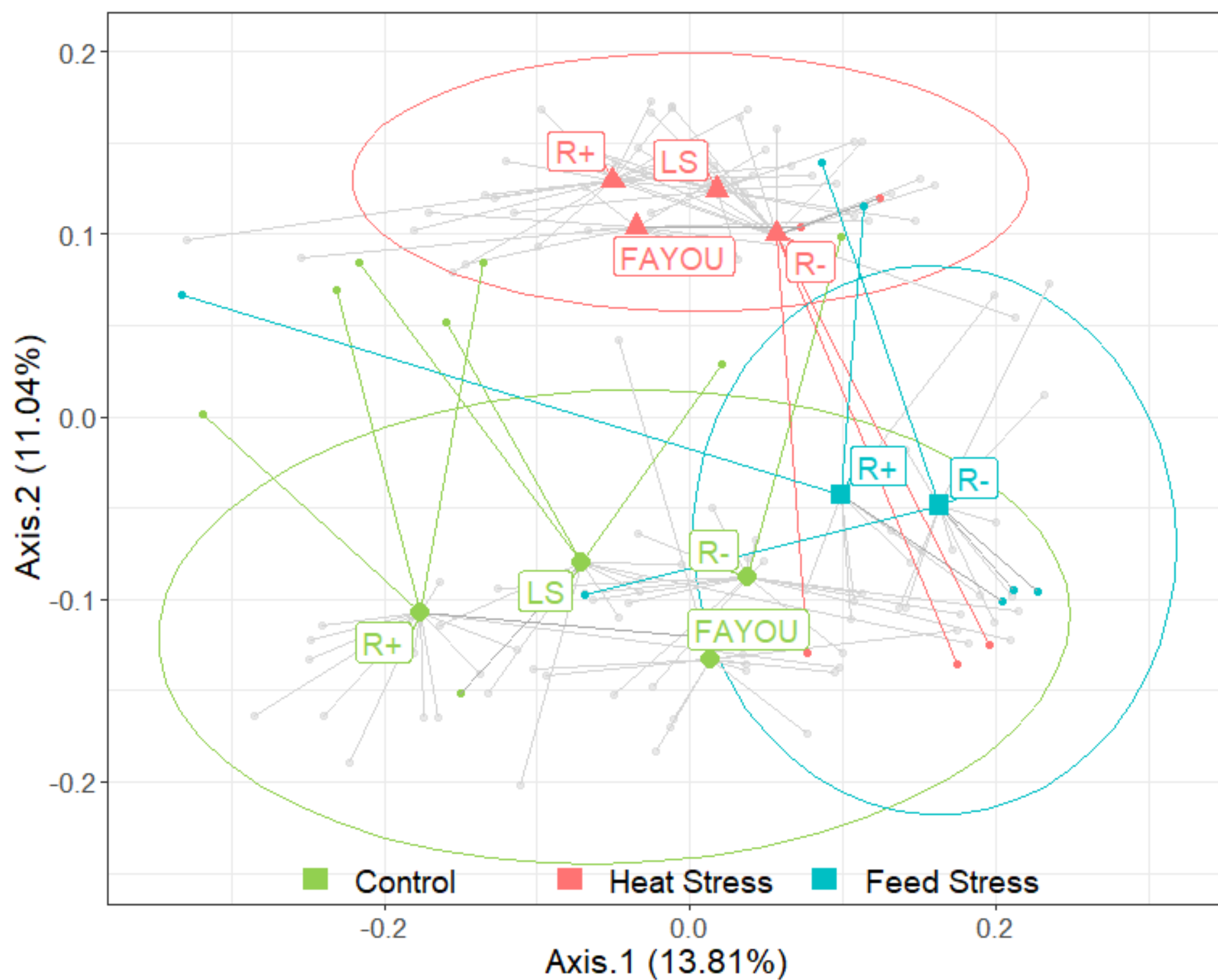
- Validate that sequencing is quite well reproducible even if the composition profile still shows variability, and confirm the high natural variability
- Identify 6 outliers samples that need to be removed

➔ Next step is to select one sample per animal by keeping the closest sample to the average profile of its group or with the deepest sequencing.

REPLICATE'S BEHAVIOR : clustering using Jaccard distance

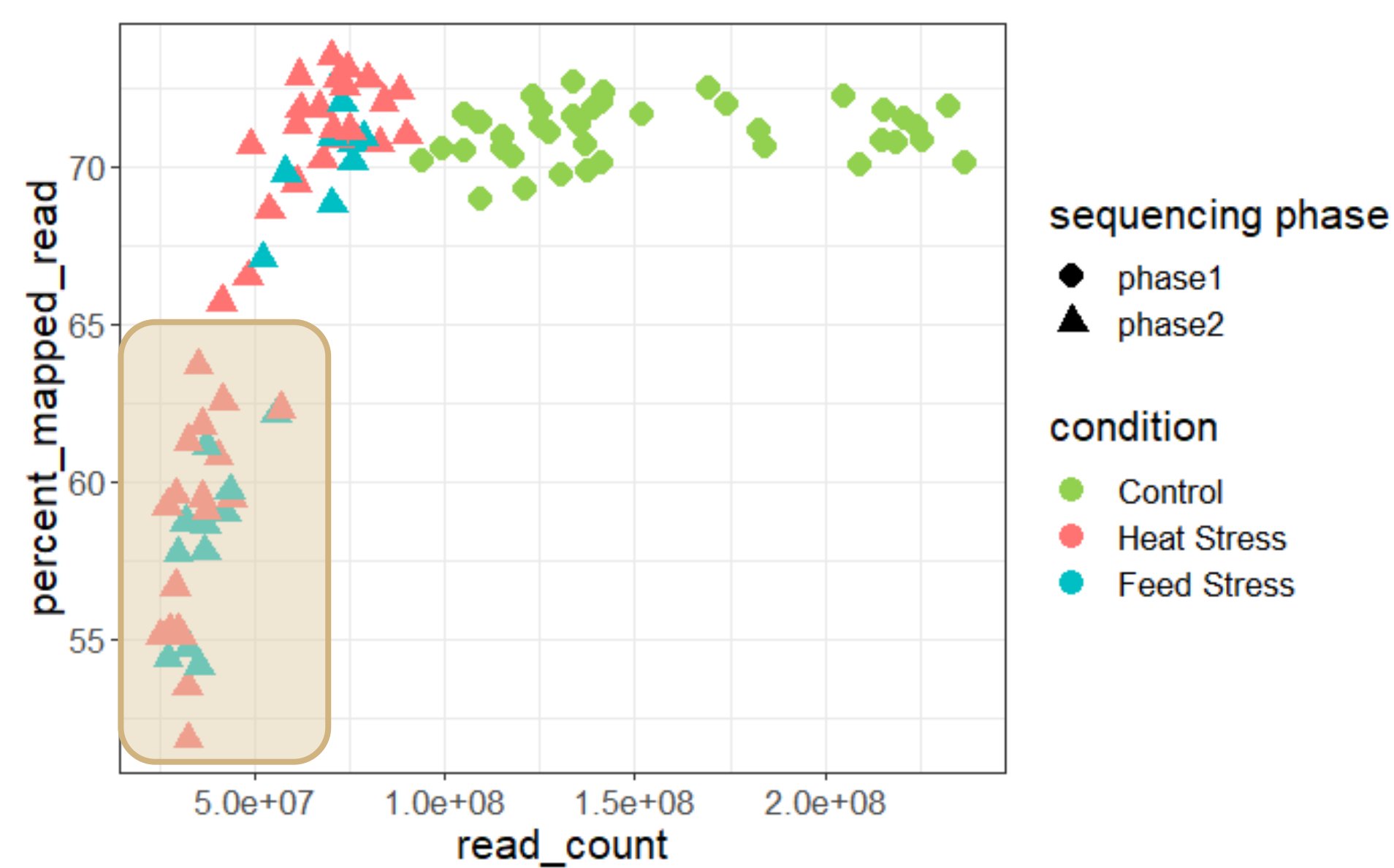


OBSERVATION : MDS on Jaccard distance



How to choose the right protocols for metagenomics shotgun and metatranscriptomics ?

Metagenomic samples were produced in two batches. Sequences have been quality checked and aligned on the internal gene chicken gut catalog using meteor (3).



➔ Is the low mapping rate due to biological reason or technical bias ?

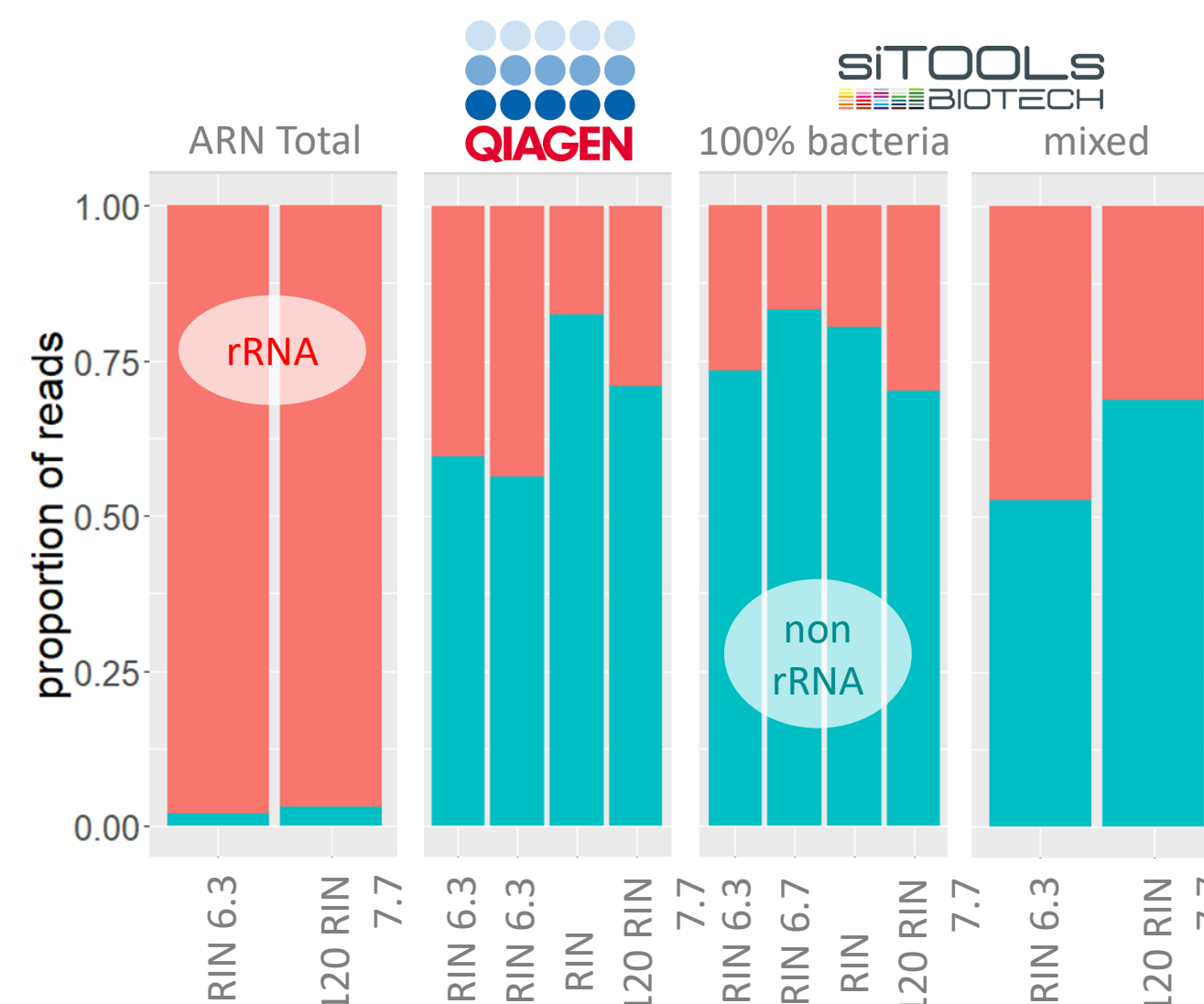
- No obvious biological reasons were discovered by unmapped read assembly and taxonomical affiliation.
- History of sample experiments highlighted a technical bias in the second phase DNA extraction protocol.

➔ Reextraction and resequencing of all stress samples as done in phase1 are in progress.

Metatranscriptomics is still an emerging omics in gut microbiota study, with a delicate protocol step, the ribodepletion.

➔ Which ribodepletion kit is most suitable for our chicken caecum samples?

Remaining rRNA were identified through SortMeRNA (4), kits were evaluated on their ease of use, cost, efficiency, sensitivity, and reproducibility.



Setting up a ribodepletion kit test project :

- 3 kits : Illumina, Qiagen, SiTools (2 probe mixes)
- 3 samples of different RIN Scores
- 1 constraint : ribodepletion experiment at Jouy en Josas and sequencing at getPlage Toulouse.

- Illumina : unusable in our experimental context, very expensive
 - SiTools (100% bacteria) vs QIAGEN :
 - equivalent cost and efficiency
 - probably a better reproducibility for QIAGEN but a higher RIN score sensitivity
 - in our experimental context, a greater ease of use for SiTools
- ➔ SiTools 100% bacteria was retained and sequences on the project samples were just received.

METAGENOMICS

METATRANSCRIPTOMICS

References

- 1 : F.Escudié et al., 2018, FROGS: Find, Rapidly, OTUs with Galaxy Solution, Bioinformatics, Volume 34, Issue 8, Pages 1287–1294
- 2 : McMurdie and Holmes, 2013, phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLoS ONE. 8(4):e61217
- 3 : Pons N et al., 2010, METEOR, a platform for quantitative metagenomic profiling of complex ecosystems. JOBIM 2010 conference.
- 4 : Evguenia K. et al. 2012, SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data, Bioinformatics, Volume 28, Issue 24, Pages 3211–3217

Fundings

ANR project Chickstress, for animal production; European project Feed-a-Gene, for metabarcoding sequencing; France Génomique project MetaChick for metagenomic sequencing and INRAE AG department project ChickMetaT for metatranscriptomic.

Acknowledgements

The authors are grateful to the Genotoul bioinformatics platform Occitanie (doi: 10.15454/1.5572369328961167E12) and Migale bioinformatics facility (doi: 10.15454/1.5572390655343293E12) for providing help, computing and storage resources. Samples have been extracted by INRAE @Bridge and SAMBO platforms, and sequences have been produced by INRAE Get-Plage and Genoscope platforms.

